

für A 1 im

2003 P 10347



34

(19) BUNDESREPUBLIK

DEUTSCHLAND



DEUTSCHES
PATENT- UND
MARKENAMT

(12) **Offenlegungsschrift**
(10) **DE 100 09 761 A 1**

(51) Int. Cl. 7:

G 06 K 9/36

(21) Aktenzeichen: 100 09 761.8
(22) Anmeldetag: 1. 3. 2000
(43) Offenlegungstag: 20. 9. 2001

(71) Anmelder:

CGK Computer Gesellschaft Konstanz mbH, 78467
Konstanz, DE

(74) Vertreter:

Zedlitz, P., Dipl.-Inf.Univ., Pat.-Anw., 80331
München

(72) Erfinder:

Kochert, Wilfried, Dipl.-Ing. (FH), 78467 Konstanz,
DE

(56) Entgegenhaltungen:

US 59 99 666 A
US 52 50 804 A

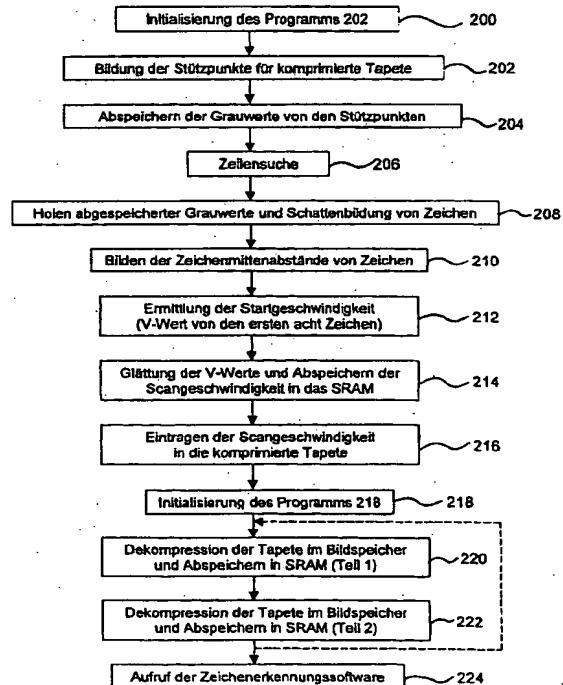
HIROMICHI FUJISAWA et.al "Segmentation
Methods
for Character Recognition: From Segmentation to
Document Structure Analyses" Proc. of the IEEE,
Vol. 80, No. 7, July 1992, S. 1079-1092;

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

Prüfungsantrag gem. § 44 PatG ist gestellt

(54) Bildaufnahmeverfahren und -Vorrichtung

(57) Die Erfindung betrifft ein Bildaufnahmeverfahren und eine Bildaufnahmeverrichtung, insbesondere einen handgeführten Scanner, um abgetastete Bilddaten verzerrungsfrei einer OCR-Software zuzuführen. Die Bildaufnahmeverrichtung enthält eine Aufnahmeverrichtung zum Abtasten einer Zeichenzeile und eine Verarbeitungsvorrichtung. Die Verarbeitungsvorrichtung bestimmt anhand der abgetasteten Bilddaten die Zeichenteilung der abgetasteten Zeichenzeile. Dies kann zum einen anhand der Zeichenabstände geschehen. Alternativ dazu kann die Übergangsschärfe der Zeichenkanten verwendet werden. Die Zeichenteilung wird anhand der Zeichenabstände ermittelt und mit der Übergangsschärfe in Zweifelsfällen überprüft.



DE 100 09 761 A 1

Beschreibung

- Die Erfindung betrifft ein Bildaufnahmeverfahren und eine Bildaufnahmeverrichtung zur Aufnahme von Bilddaten einer Zeichenzeile, wobei eine Aufnahmeverrichtung zur Abtastung einer Vorlage, um die Bilddaten einer Zeichenzeile zu lesen, und einer Verarbeitungsvorrichtung vorgesehen sind, um unter Berücksichtigung einer tatsächlichen Zeichenteilung die Bilddaten für eine OCR-Erkennung vorzubereiten, um so verzerrungsfreie Bilddaten zu erhalten.
- In den letzten Jahren ist die OCR-Erkennung, also die Umwandlung von Bilddaten in Textdaten, zunehmend wichtiger geworden. Voraussetzung für eine einwandfreie OCR-Erkennung ist es allerdings, dass die Bilddaten der OCR-Software mit einer gleichbleibend guten Qualität zugeführt werden.
- Wenn die Bilddaten für die OCR-Erkennung über eine Bildaufnahmeverrichtung, beispielsweise einen Scanner, eingelesen werden, stellt sich das Problem, dass die erhaltenen Bilddaten aufgrund von Scannereigenschaften nicht verzerrungsfrei sind, was zu Schwierigkeiten bei der OCR-Erkennung führen kann. Dies ist insbesondere bei handgeführten Scanern ein ernstes Problem.
- Bei einem handgeführten Scanner werden die Zeichen je nach Geschwindigkeit beim Scannen, die zwischen 5 und 150 cm/sec. liegen kann, beliebig stark verzerrt. Um dieses Problem zu lösen, gab es bisher zwei Ansätze. Zum einen offenbart US-5,250,804 einen handgehaltenen Scanner mit einem Taktrad. Wenn der Scanner über eine Vorlage bewegt wird, dreht sich das Taktrad und gibt ein Maß für die Geschwindigkeit, das an den Computer übermittelt wird. Ein Nachteil dieser Bauweise ist es, dass nur sehr geringe Scangeschwindigkeiten möglich sind, eine sehr saubere Scanneführung erforderlich ist, und dass das Taktrad eine sehr begrenzte Lebensdauer hat.
- Ein zweiter Ansatz sieht vor, ein zweidimensionales Photodiodenarray in der Lesevorrichtung einzusetzen. Wird dieses Photodiodenarray hinreichend groß gewählt, ist es möglich, ein Normzeichen mit einem Scannzyklus zu erfassen. Der Nachteil hierbei ist, dass zweidimensionale Arrays relativ teuer sind, und dass breite Normschriften nicht gelesen und erkannt werden können, da Teile eines Zeichens vom Flächensor nicht erfasst werden.
- Aufgabe der Erfindung ist es, ein verbessertes Bildaufnahmeverfahren und eine verbesserte Bildaufnahmeverrichtung zu schaffen, die weitgehend verzerrungsfrei arbeiten.
- Erfindungsgemäß wird die Aufgabe durch ein Verfahren nach Anspruch 1 bzw. eine Vorrichtung nach Anspruch 11 gelöst. Die abhängigen Ansprüche betreffen weitere vorteilhafte Aspekte der Erfindung.
- Bei dem erfindungsgemäßen Verfahren zur Aufnahme von Bilddaten einer Zeichenzeile wird eine Zeichenzeile einer Vorlage zum Lesen von Bilddaten abgetastet und die Bilddaten für eine OCR-Erkennung vorbereitet, wobei die Zeichenteilung beruhend auf den Bilddaten bestimmt wird.
- Bei einem solchen Verfahren ist die Verwendung eines Taktrades, welches relativ schnell kaputt gehen kann, entbehrlich. Zudem lassen sich Scangeschwindigkeiten von mehr als 1 m/Sek. erreichen. Die Anforderungen an die Genauigkeit der Scanneführung sind geringer und das Verfahren ist einfacher durchführbar, da keinerlei bewegliche Teile verwendet werden.
- Eine vorteilhafte Ausgestaltung des erfindungsgemäßen Verfahrens ist dadurch gekennzeichnet, dass die Zeichenteilung aus dem Abstand aufeinanderfolgender Zeichen bestimmt wird. Hierdurch wird in einfacher Weise ein jeweils aktuelles Maß für die Zeichenteilung erhalten.
- Eine vorteilhafte Ausgestaltung des erfindungsgemäßen Verfahrens ist dadurch gekennzeichnet, dass zusätzlich die Übergangsschärfe an den Zeichenkanten bestimmt wird, wobei die Übergangsschärfe zur Überprüfung der aus dem Abstand zwischen aufeinanderfolgenden Zeichen ermittelten Zeichenteilung dient. Hierdurch lässt sich die Zuverlässigkeit der ermittelten Zeichenteilung erhöhen, was insbesondere dann von Bedeutung ist, wenn die Zeichen in horizontaler Richtung Unterbrechung aufweisen, so dass irrelevante Zeichenzwischenräume entstehen, die zu einer fehlerhaften Ermittlung der Zeichenteilung führen könnten.
- Eine vorteilhafte Ausgestaltung des erfindungsgemäßen Verfahrens ist dadurch gekennzeichnet, dass die gelesenen Bilddaten komprimiert werden. Da insbesondere bei langsamem Scannen sehr große Datenmengen entstehen können, lässt sich auf diese Weise ein wesentliches Maß an Speicherplatz und -zeit einsparen.
- Eine vorteilhafte Ausgestaltung des erfindungsgemäßen Verfahrens ist dadurch gekennzeichnet, dass die komprimierten Bilddaten als die Grauwerte in einem SRAM einer Zeichenerkennungseinheit für Grauwert-Videodaten abgespeichert werden. Das Abspeichern in Form von Grauwerten hat gegenüber dem Abspeichern in binarisierter Form den Vorteil, dass sich durch die Analyse von Weiß-schwarz- beziehungsweise Schwarz-weiß-Übergängen relevante Zeichenteile von Zeichen näher untersuchen lassen.
- Die erfindungsgemäße Bildaufnahmeverrichtung zur Aufnahme von Bilddaten einer Zeichenzeile enthält eine Aufnahmeverrichtung zur Abtastung einer Vorlage, um die Bilddaten einer Zeichenzeile zu lesen und eine Verarbeitungsvorrichtung, um unter Berücksichtigung einer tatsächlichen Zeichenteilung die Bilddaten für eine OCR-Erkennung vorzubereiten, um so verzerrungsfreie Bilddaten zu erhalten. Die erfindungsgemäße Vorrichtung ist dadurch gekennzeichnet, dass die Verarbeitungsvorrichtung ausgestattet ist, um den Abstand zwischen aufeinanderfolgenden Zeichen der Zeichenzeile und somit ein Maß für die tatsächliche Zeichenzeile zu bestimmen.
- Ein Vorteil der Erfindung besteht darin, dass kein Taktrad nötig ist, somit also keine mechanischen Teile vorliegen, die relativ schnell kaputtgehen können. Mit der erfindungsgemäßen Vorrichtung sind Scangeschwindigkeiten von mehr als 1 m/sec. möglich. Auch die Anforderungen an die Genauigkeit der Scanneführung sind geringer als beim Stand der Technik. Die gesamte Vorrichtung wird durch das Vermeiden mechanischer Teile billiger, robuster und einfacher in der Handhabung.
- Eine vorteilhafte Ausgestaltung der erfindungsgemäßen Bildaufnahmeverrichtung ist dadurch gekennzeichnet, dass die Verarbeitungsvorrichtung in der Lage ist, den Abstand zwischen aufeinanderfolgenden Zeichen zu ermitteln und aus diesem Abstand die Zeichenteilung bestimmen.
- Eine vorteilhafte Ausgestaltung der erfindungsgemäßen Bildaufnahmeverrichtung ist dadurch gekennzeichnet, dass die Verarbeitungsvorrichtung derart ausgestaltet ist, dass sie zusätzlich die Übergangsschärfe an den Zeichenkanten bestimmt, wobei die Übergangsschärfe ein Maß für die Zeichenteilung ist. In Abhängigkeit der Geschwindigkeit des

DE 100 09 761 A 1

Scannvorgangs gibt es an Schwarz-Weiß- bzw. Weiß-Schwarz-Übergängen sogenannte Graukegel. Die Größe dieser Graukegel ist ebenfalls ein Maß für die Scangeschwindigkeit bzw. für die Zeichenteilung. Insbesondere bei schlechten Vorlagen, bei denen einzelne Zeichen der abzuscannenden Zeichenzeile unvollständig wiedergegeben sind, bietet die Bestimmung der Übergangsschärfe eine zusätzliche Überprüfungsmöglichkeit für den vorhergehend bestimmten Zeichenabstand. Die Vorrichtung wird auf diese Art noch leichter in der Handhabung.

Eine vorteilhafte Ausgestaltung der erfundungsgemäßen Bildaufnahmeverrichtung ist dadurch gekennzeichnet, dass die Verarbeitungsvorrichtung zusätzlich die Übergangsschärfe an den Zeichenkanten bestimmt, wobei die Übergangsschärfe zur Überprüfung der aus dem Zeichenabstand ermittelten Zeichenteilung dient.

Eine vorteilhafte Ausgestaltung der erfundungsgemäßen Bildaufnahmeverrichtung ist dadurch gekennzeichnet, dass die Aufnahmeverrichtung als handgeführter Scanner ausgeschaltet ist. Von diesem handgeführten Scanner werden die gewonnenen Bilddaten durch ein geeignetes Mittel, beispielsweise ein Kabel oder eine Infrarotschnittstelle, an die Verarbeitungsvorrichtung der Bildaufnahmeverrichtung übertragen. Insbesondere bei handgeführten Scannern variiert die Scangeschwindigkeit sehr stark. Um dennoch eine effiziente OCR-Erkennung durchführen zu können, ist daher eine zuverlässige Gewinnung verzerrungsfreier Daten erforderlich.

Eine vorteilhafte Ausgestaltung der erfundungsgemäßen Bildaufnahmeverrichtung ist dadurch gekennzeichnet, dass die Vorrichtung einen Speicher um die Bilddaten zwischenzuspeichern. In dem Speicher können daher in vorteilhafter Weise die abgetasteten Bilddaten und die entzerrten Bilddaten vor der Verarbeitung durch die OCR-Software zwischen gespeichert werden.

Eine vorteilhafte Ausgestaltung des erfundungsgemäßen Verfahrens ist dadurch gekennzeichnet, dass es das Komprimieren der eingelesenen Bilddaten und Speichern als Grauwerte in einem SRAM-Puffer, das Abspeichern von Grauwertstützpunkten zum Bilden einer komprimierten Tapete, das Bilden von Zeichenmittensabständen aufeinanderfolgender Zeichen, das Bilden von Scangeschwindigkeitswerten und Eintragen dieser Scangeschwindigkeitswerte in die komprimierte Tapete und das Dekomprimieren der komprimierten Tapete umfasst. Auf diese Weise lässt sich ein geschwindigkeitskorrigiertes Band von Zeichen erhalten, welches anschließend einer OCR-Zeichenerkennung zugeführt werden kann.

Eine vorteilhafte Ausgestaltung des erfundungsgemäßen Verfahrens ist dadurch gekennzeichnet, dass zum Bilden der Grauwertstützpunkte die als Grauwerte abgespeicherten Bilddaten zeilenweise unter Bildung von Grauwertdifferenzen zu einem für jede Zeile ermittelten Weißwert untersucht werden, wobei bei Überschreiten einer vorbestimmten Grauwertdifferenz ein Grauwertstützpunkt in den SRAM-Puffer eingetragen wird. Anhand der auf diese Weise ermittelten Grauwertstützpunkte lassen sich die zu einer Zeile gehörenden relevanten Pixel in einfacher Weise durch Addition der Differenz der Grauwertstützpunkte zwischen minimalem und maximalem Grauwert anhand einer dynamisch einstellbaren Schwellen ermitteln.

Eine vorteilhafte Ausgestaltung des erfundungsgemäßen Verfahrens ist dadurch gekennzeichnet, dass vor dem Bilden von Zeichenmittensabständen aufeinanderfolgender Zeichen Grauwerte unterhalb eines jeweiligen Schwellenwertes auf 0 gesetzt werden, wobei die Schwellenwerte jeweils zeilenweise abhängig von einem maximalen, in horizontaler Richtung ermittelten Grauwert gebildet werden, wodurch sich eine Schattenbildung der Zeichen durchführen lässt.

Eine vorteilhafte Ausgestaltung des erfundungsgemäßen Verfahrens ist dadurch gekennzeichnet, dass die Scangeschwindigkeitswerte aus den Mittenabständen aufeinanderfolgender Zeichen mittels einer Mehrheitssuche bestimmt werden. Durch die Mehrheitssuche lassen sich Verfälschungen aufgrund von Unterbrechungen im Zeichen oder bei Zeichen ohne Weißlücke vermeiden.

Eine vorteilhafte Ausgestaltung des erfundungsgemäßen Verfahrens ist dadurch gekennzeichnet, dass bei der Ermittlung der Scangeschwindigkeiten die Steigungsmaße der Schwarz-weiß-beziehungsweise Weiß-schwarz-Übergänge der Zeichenteile in horizontaler Richtung ermittelt werden. Hierdurch werden Störungen bei der Ermittlung der Zeichenteilung bei verstümmelten Zeichen mit irrelevanten Zeichenzwischenräumen vermieden.

Ausführungsbeispiele der Erfindung ergeben sich aus der vorliegenden Beschreibung der beiliegenden Zeichnungen, in denen zeigt:

Fig. 1 Rohbilddaten nach der Abtastung mit einem handgeführten Scanner;

Fig. 2 die entzerrten Bilddaten;

Fig. 3 die Grauwertverteilung der Zeile 28 aus Fig. 1;

Fig. 4 die Grauwertverteilung der Zeilen 27, 28 und 29 eines Ausschnittes aus Fig. 1;

Fig. 5 eine Darstellung der Grauwertstützpunktverteilung;

Fig. 6 ein Flussdiagramm des erfundungsgemäßen Verfahrens.

Im Folgenden wird die Erfindung detailliert anhand einer bevorzugten Ausführungsform beschrieben.

Erfundungsgemäß werden OCR-Zeichen optisch abgetastet, die so erhaltenen Videodaten komprimiert, die Zwischenräume der OCR-Zeichen aufgefunden und der Abstand zweier Zeichen ermittelt, um dadurch ein Maß für die aktuelle Zeichenteilung zu bekommen. Die Kompression der Videodaten ist nicht nötig, aber sinnvoll, da bei langsamem Scannen sehr große Datenmengen entstehen können und damit viel Speicher und Rechenzeit benötigt wird. Bei der Dekomprimierung der Videodaten wird die Zeichenteilung strichgenau berücksichtigt. Die Soll-Zeichenteilung für gängige OCR-Normschriften beträgt im Normalfall zehn Zeichen pro Zoll. Auf diese Zeichenteilung werden die Zeichen normiert.

In manchen Fällen kann es bei Zeichen vorkommen, dass sie in horizontaler Richtung Unterbrechungen haben, so dass ein irrelevanter Zeichenzwischenraum erkannt wird. Diese Zwischenräume können zu einer falschen Zeichenteilung führen. Aus diesem Grund wird in der erfundungsgemäßen Ausführungsform ein zweiter Algorithmus angewendet, der die Scangeschwindigkeit von Zeichen näherungsweise ermittelt. Das Ergebnis dient dazu, herauszufinden, ob bzw. welche Zeichenteilung plausibel ist.

Der zweite Algorithmus betrachtet den Weiß-Schwarz- bzw. den Schwarz-Weiß-Übergang einer Linie eines Zeichens in horizontaler Richtung. In der Zeichenzeile werden relevante Zeichenteile von Zeichen gesucht. Relevant sind Zeichenteile, wenn sie größer als zwei Zeichenpixel breit sind und von einem Zeichen mit einer annähernd senkrechten Zeichenkante stammen. Hierfür werden die Nachbarpixel in vertikaler Richtung untersucht. Wenn ein derart relevanter

DE 100 09 761 A 1

Zeichenteil vorliegt, kann die Länge des "Graukeils" am Weiß-Schwarz- bzw. Schwarz-Weiß-Übergang, nämlich die Unschärfe der Kante, ein Maß für die Abtastgeschwindigkeit geben. Hierfür ist es erforderlich, dass die Videodaten nicht binarisiert, sondern Grauwerte enthalten.

Am Ende entsteht eine geschwindigkeitskorrigierte und somit entzerrte Zeichenzeile. Diese Zeichenzeile kann nun einer OCR-Zwischenerkennung zugeführt werden, die die Bilddaten in Textdaten umwandelt.

Entsprechend einer bevorzugten Ausführungsform führt die erfindungsgemäße Bildaufnahmeverrichtung folgende Verarbeitungsprozedur durch. Zunächst wird mit einer Aufnahmeverrichtung eine Zeichenzeile abgescannt. Das Ergebnis dieses Scannvorgangs ist in Fig. 1 gezeigt. Durch die unterschiedlich schnelle Bewegung der Aufnahmeverrichtung über die Zeichenzeile entsteht eine Verzerrung der einzelnen Zeichen. Diese Verzerrung erschwert es den gängigen OCR-Programmen die abgetastete Zeichenzeile in Text umzuwandeln.

Das erfindungsgemäße Verfahren zur Aufnahme von Bilddaten wird im Folgenden anhand des in Fig. 6 dargestellten Flussdiagramms detailliert erläutert.

Mittels der in Fig. 6 dargestellten Programmabfolge werden die eingelesenen Videodaten komprimiert und in das SRAM einer Zeichenerkennungseinheit für Grauwert-Videodaten mit 6 Bit Grauwerten abgespeichert. Nachdem der Beleg gescannt wurde, werden die komprimierten Daten dekomprimiert und geschwindigkeitskorrigiert in den Bildspeicher abgespeichert und von dort wieder in das SRAM als sogenannte Tapete abgespeichert. Nachdem die Tapete wieder in den Bildspeicher übertragen wurde, kann das Programm zur Zeichenerkennung durchgeführt werden.

Gemäß Fig. 6 läuft das beim handgeföhrten Lesen der Bilddaten verwendete Programm folgendermaßen ab:

Zunächst wird mittels eines Programms 200 ein Programm 202 zur Bildung von Stützpunkten für die komprimierte Tapete initialisiert. Bei Durchführung des Programms 200 wird ein SRAM-Puffer aufgebaut. In diesem Puffer werden die Daten von 64 Zeilen abgespeichert. Pro Zeile gibt es vier Werte: Spaltenadresse, Weißwert, letzter abgespeicherter Grauwert und Status.

Die Hardware arbeitet im Transparentmodus, das heißt, die Grauwerte des Scanners werden unverfälscht in die Zeichenerkennungseinheit für Grauwert-Videodaten eingetragen. Der Weißwert wird aus den ersten 32 Pixeln der jeweiligen Zeile gebildet, indem der weißeste Wert gesucht wird. Der Status ist am Anfang Null.

Nach dieser Initialisierung des SRAM-Puffers werden die Spaltenadressen und die Zeilenadressen der Pixel der Anfangsspalte in den Puffer eingetragen. Dies sind Pixel von 64 Zeilen. Am Ende einer Spalte wird ein Zeilenende-Code (0xc000) und die Zeilenadresse des letzten Pixels eingetragen.

Anschließend wird das Programm 202 zur Bildung der Stützpunkte für die komprimierte Tapete durchgeführt. Hierbei werden pro Spalte die Pixel von Zeile 0 bis Zeile 63 untersucht, ob sie in horizontaler Richtung rechts Grauwertdifferenzen von größer als 3 in Bezug auf den Weißwert aufweisen. Ist das nicht der Fall, wird das nächste Pixel (die nächste Zeile) in horizontaler Richtung untersucht.

Ist die jeweilige Grauwertdifferenz größer als 3, wird in der vorherigen Zeile ($n - 1$) der Status 1 gesetzt. Mit der nächsten Spalte wird dieses Pixel (Grauwertstützpunkt) dann in den Speicher SRAMZ mit der Spaltenadresse und Zeilenadresse abgespeichert. In der entsprechenden Zeile im Puffer WEZ wird der Status 8 eingetragen. Diese Zeile wird nun spaltenweise verfolgt. In den SRAMZ-Puffer wird immer dann ein Pixel mit Spaltenadresse und Zeilenadresse übertragen, wenn die Grauwertdifferenz größer als 3 wird oder wenn ein Umkehrpunkt vorhanden ist (weiß, schwarz, weiß oder schwarz, weiß, schwarz). Erreicht ein Grauwert bis auf 3 Grauwertstufen den Weißwert, wird der Status auf 4 gesetzt. Mit der nächsten Spalte wird das Pixel nach SRAMZ übertragen und der Status in dieser Zeile wird auf 0 gesetzt. Jetzt wird eine neue Grauwertdifferenz gesucht.

So wie mit dieser Zeile wird mit allen Zeilen verfahren.

In Zeile 63 ist das Spaltenende erreicht. In den Speicher SRAMZ wird der Zeilenende-Code 0xc000 eingetragen. Es werden so viele Spalten abgearbeitet, wie im Bildspeicher vorhanden sind. Die Bilddaten werden vom Programm 204 in das SRAM eingetragen. Liegen keine Bilddaten mehr vor, wird das Programm 202 verlassen und das Programm 204 aufgerufen.

Das Programm 204 dient zum Abspeichern der Grauwerte von den Stützpunkten. Hierbei werden die Adressen der Pixel aus dem Puffer SRAMZ geholt und die Pixel im Bildspeicher adressiert. Die ausgelesenen Grauwerte (Byte) werden wortweise verpackt und in den Puffer WEZ (gleicher Name, neuer Puffer) in das SRAM eingetragen. Am Anfang einer Spalte wird die Spaltenadresse und am Ende einer Spalte der Zeilenende-Code (0xc000) ins SRAM eingetragen.

Sind alle bisher gescannten Pixel im SRAM, wird untersucht, ob der Beleg noch vorliegt (Belegimpulsbildung). Ist dies der Fall, wird die Anzahl der Spalten berechnet, die der Einschreibzähler inzwischen in den Bildspeicher geschrieben hat. Diese Anzahl wird dem Programm 202 mitgeteilt, welches wieder aufgerufen wird.

Liegt kein Beleg mehr vor, wird die Schleife verlassen, und das Programm 206 wird aufgerufen. In dem Puffer WEZ liegen jetzt die komprimierten Bilddaten.

Das Programm 206 dient zum Suchen der gedruckten Zeile innerhalb der gesamten Tapete. Es ermittelt die Oberkante und die Unterkante der gesamten Zeile. Eingangsdaten sind die komprimierten Bilddaten aus dem Puffer WEZ. Das Programm 206 findet die relevanten Pixel, die zu einer Zeile gehören. Es addiert in jeder Zeile die Differenz der Grauwert-Stützpunkte zwischen minimalem und maximalem Grauwert. Mit einer dynamisch einstellbaren Schwelle wird am Ende die obere und untere Kante der Zeile gefunden.

Das Programm 208 dient zum Holen abgespeicherter Grauwerte und zur Schattenbildung von Zeichen. Hierbei wird der vom Programm 206 ausgegebene Bereich ermittelt, pro Spalte der maximale Grauwert bestimmt und für diese Spalte abgespeichert. Hierbei handelt es sich um die Grauwerte der komprimierten Bilddaten. Parallel dazu wird pro Zeile eine Schwelle gebildet, die abhängig von einem maximalen Grauwert ist, der in horizontaler Richtung gebildet wird. Diese Schwelle wird nach jeder Spalte in jeder Zeile um zwei Grauwertstufen dekrementiert (schwarz = 3f, weiß = 0). Mit der nächsten Spalte kann ein neuer Grauwert (Grauwertstützpunkt) in dieser Zeile vorhanden sein. Der neue wird mit dem alten verglichen, der größere wird abgespeichert. Der Maximalwert der Schwelle beträgt 15, der Minimalwert 4. Liegt der Grauwert unterhalb der ermittelten Schwelle, wird er auf 0 gesetzt. Diese neuen Grauwerte werden in das SRAM geschrieben (neuer Puffer REZ).

DE 100 09 761 A 1

Anschießend wird das Programm 210 durchgeführt, mittels dem die Zeichenmittenstände von Zeichen ermittelt werden. Aus dem SRAM-Puffer (REZ) werden die Spaltenadressen und maximalen Grauwerte ausgelesen und entsprechende Segmente gebildet. Je nach Größe des maximalen Grauwerts eines Zeichens (Zeichensegments) werden bis zu 3 Weißspalten zwischen den Zeichensegmenten zugelassen, so dass ein Zeichen aus mehreren Segmenten bestehen kann. Bei sehr schwach gedruckten Zeichen kommen häufig Unterbrechungen innerhalb des Zeichens vor. Von dem gegebenenfalls zusammengesetzten Zeichen wird die Zeichenmitte berechnet. Die Zeichenmittenadresse und die Zeichenbreite werden in den SRAM-Puffer (WEZ) eingetragen. Das Dateiende wird mit 0xffff gekennzeichnet.

Anschießend wird aus den Daten des Puffers (WEZ) der Zeichenabstand zweier Zeichen ermittelt und an die Stelle der Zeichenbreite wieder in den Puffer (WEZ) eingetragen. Zuvor wird noch überprüft, ob eine größere Weißlücke zwischen den Zeichen besteht. Hierzu wird von dem Zeichenabstand die vierfache Zeichenbreite subtrahiert. Ist der verbleibende Abstand größer als 32, liegt eine größere Weißlücke vor. Im Zeichenabstand (später = V-Wert) wird das höchste Bit 1 gesetzt. Der V-Wert ist ein Maß für die Geschwindigkeit des Belegs. Ist der V-Wert 1,0, liegt die Normgeschwindigkeit vor. Das es sich um handgeführtes Lesen handelt, kann die Beleggeschwindigkeit variieren. Wenn also eine größere Zeichenlücke vorhanden ist, muss nach dieser Lücke die Geschwindigkeit neu berechnet werden. Dies wird mittels der Programme 212 und 214 durchgeführt, die dieses Bit abfragt.

Das Programm 212 ermittelt die V-Werte (a und b) für die ersten maximal 8 Werte. Der V-Wert (a) entspricht genau dem Mittenabstand zweier benachbarter Zeichen. Ist dieser Abstand 16 Spalten groß, dann liegt die Normgeschwindigkeit vor. Bei 10 Zeichen/Zoll und 160 Micrometer Pixelauflösung beträgt der Zeichenabstand 16 Spalten. Parallel dazu wird das höchste Bit des V-Wertes abgefragt. Ist dieses 1, wird die Bearbeitung beendet und Programm 214 wird aufgerufen. Es wird erst bei dem nächsten Aufruf dieses Programms 212 fortgefahrene. Die Programme 212 und 214 werden in einer Schleife solange aufgerufen, bis der Endcode 0xffff kommt.

Die V-Werte (a) können durch Unterbrechungen im Zeichen oder bei Zeichen ohne Weißlücke stark verfälscht sein. Deshalb wird eine Mehrheitssuche unter den ersten 8 Werten durchgeführt. Die Zeichenabstände, die am häufigsten vorkommen, bestimmen den V-Wert der anderen. Die Häufigkeitsverteilung wird mittels Histogramm ermittelt. Die ermittelten V-Werte werden in den Puffer (WEZ) eingetragen.

In Fällen, bei denen Zeichen verstümmelt sind, kann es zu erheblichen Störungen bei der Ermittlung der Zeichencentierung kommen, so dass keine eindeutige Entscheidung für einen V-Wert getroffen werden kann. Fehlt zum Beispiel bei einer "0" oben und unten der Querstrich, dann ist der V-Wert für dieses Zeichen um den Faktor 2 zu klein. Bei etwa Normgeschwindigkeit kann der V-Wert mit Hilfe der Plausibilität korrigiert werden. Wird aber erheblich langsamer als Normgeschwindigkeit gescannt, dann kann auf diese Art der V-Wert nicht mehr korrigiert werden. In diesem Falle werden die Weiß-schwarz- bzw. Schwarz-weiß-Übergänge des Zeichens näher untersucht. Es werden drei unmittelbar übereinander liegende Pixelzeilen in der entsprechenden Spalte betrachtet (siehe Fig. 4). Haben die drei Pixelzeilen annähernd den gleichen Kurvenverlauf und ist die Grauwertdifferenz zwischen dem hellsten und dunkelsten Pixel größer als 15, dann wird ein Steigungsmaß bei jedem Weiß-schwarz-/Schwarz-weiß-Übergang von jeder Zeile ermittelt.

Dies wird folgendermaßen durchgeführt.

In Fig. 1 sind die original gescannnten Grauwertbilddaten abgebildet. Man sieht, dass die Zeile am Anfang erheblich langsamer als am Ende des Belegs gescannt wurde.

Gemäß Fig. 3 wurde beispielhaft eine Zeile (Zeile 28) herausgegriffen. Die Bilddaten sind so aufgebaut, dass links die Spalte 0 und oben die Zeile 0 ist. Gemäß Fig. 3 ist von der Zeile 28 die Grauwertverteilung dargestellt. Man erkennt, wie sich die Steilheit der Flanke abhängig von der Scanngeschwindigkeit verändert. Als crstcs wird der Beginn einer Flanke gesucht. Es wird der Weißwert des Belegs ermittelt. Mittels einer Schwelle kann das Pixel gefunden werden, das über dieser Schwelle liegt. Jetzt wird von Pixel zu Pixel die Differenz (d) berechnet und gespeichert, bis die Differenz wieder einen Schwellenwert unterschreitet oder die Polarität wechselt. Aus den Differenzen D1 bis Dn wird der maximale Wert dmax ermittelt. Nun wird ein Wert f nach folgender Formel berechnet.

$$f = \text{dmax} \left(\frac{1}{d_1} + \frac{1}{d_2} + \dots + \frac{1}{d_n} \right)$$

Der Wert wird nun über eine Tabelle korrigiert, und man erhält den V-Wert (b).

Beispiel

Zeichen 1: f = 3,4 Korrekturtabelle ergibt 2,37 = V-Wert (b).

Zeichen 2: f = 2,3 Korrekturtabelle ergibt 1,50 = V-Wert (b).

Zeichen 3: f = 1,5 Korrekturtabelle ergibt 1,00 = V-Wert (b) = Normgeschwindigkeit.

Hierbei bedeutet zum Beispiel ein V-Wert von 2,37, dass das Zeichen um diesen Faktor in horizontaler Richtung verkleinert werden muss. Bei einem V-Wert = 1 wird das Zeichen nicht verändert.

Anschießend wird noch eine Wertbildung der V-Werte von drei unmittelbar untereinander liegenden Zeilen durchgeführt. Dieser V-Wert (b) ist ein Maß für die Scanngeschwindigkeit. Der V-Wert (b) wird jetzt mit dem V-Wert (a) verglichen. Weicht der V-Wert (b) mehr als 50% vom V-Wert (a) ab, dann wird der V-Wert (b) für den V-Wert eingesetzt.

Anschießend wird das Programm 214 zur Glättung der V-Werte und zum Abspeichern der Scanngeschwindigkeit in das SRAM durchgeführt. Hierbei wird der Mittelwert von 4 V-Werten gebildet. Die letzten drei V-Werte (Belegende) bekommen den Wert des zuletzt ermittelten V-Wertes. Bei der Glättung wird vorher abgefragt, ob der nächste Wert einen zu großen Abstand zum vorhergehenden hat. Ist dies der Fall, wird er gleich dem vorhergehenden gesetzt, da eine zu große Abweichung der Geschwindigkeit von Zeichen zu Zeichen nicht vorkommen kann.

Ist bei einem V-Wert das höchste Bit 1, so wird dieser V-Wert noch überschrieben, und Programm 212 wird aufgerufen. Mit den Endcode 0xffff wird die Schleife verlassen. Anschließend wird Programm 216 aufgerufen.

DE 100 09 761 A 1

Mittels Programm 216 wird der ermittelte V-Wert in die "komprimierte Tapete" (Puffer REZ) eingetragen. In Puffer WEZ befindet sich die Spaltenmittadenresse der Zeichen und der V-Werte.

Dies wird folgendermaßen durchgeführt:

5 Im RAM Puffer REZ und WEZ sind jeweils die Spaltenadressen am Anfang einer Spalte abgespeichert. Die Spaltenadressen laufen syncron zueinander. Es wird nun die Spaltenadresse 1 des Puffers REZ als Startadresse genommen und die V-Wert aufaddiert. Da es sich um eine komprimierte Tapete handelt, muss nicht notwendigerweise eine Spaltenadresse 2 existieren. Ist die nächste Spaltenadresse zum Beispiel 5, dann muss der V-Wert n-mal auf die Spaltenadresse 1 addiert werden. Mit jeder Addition wird eine fortlaufende Adresse inkrementiert. Während der Addition wird die neue Spaltenadresse mit der im Puffer WEZ verglichen. Ist die neue Spaltenadresse ganzzahlig größer als die Puffer WEZ, wird die fortlaufende Adresse als neue Adresse in REZ übernommen.

10 Die Beleggeschwindigkeit kann schneller, gleich oder langsamer als die Sollgeschwindigkeit sein. Ist sie schneller, werden die Bilddaten gespreizt. Ist die Beleggeschwindigkeit lang, werden die Bilddaten gestaucht, und entspricht die Beleggeschwindigkeit der Normgeschwindigkeit, dann verändern sich die Bilddaten nicht.

15 Alle Programme 206 bis 216 wurden auf den komprimierten Bilddaten durchgeführt. Das Programm 218 dient zur Initialisierung des Programms 120, welches die komprimierte Tapete wieder dekomprimiert.

Außerdem speichert das Programm 220 die dekomprimierte Tapete im Bildspeicher ab.

20 Hierbei werden Spalte für Spalte die im Puffer REZ abgespeicherten Grauwertstützpunkte zeilenweise miteinander verbunden und im Bildspeicher abgespeichert. Hierbei werden der Spaltenabstand zwischen den Grauwertpixeln und die Höhe der Grauwerte berücksichtigt, so dass es unterschiedliche Grauwertkeile in horizontaler Richtung gibt. Auf diese Weise entsteht ein in horizontaler Richtung korrigiertes Bild im Bildspeicher, das zur Zeichenerkennung verwendet wird.

25 Das Programm 222 überträgt die Bilddaten des Bildspeichers in den FRAM Puffer (neuer WEZ). Hier liegen nach Beendigung der Bearbeitung des Scannvorgangs die endgültigen Bilddaten (siehe Fig. 2).

Da der Bildspeicher nur 64 Spalten groß ist, zur Rekonstruktion der Bilddaten aus der komprimierten Tapete aber etwa 30 Spalten benötigt werden, muss nach 32 Spalten die Hälfte des Bildspeichers in das SRAM abgespeichert werden. Das Programm 220 kann dann den Bildspeicher wieder laden.

30 Die Schleife aus den Programmen 220 und 222 wird solange durchgeführt, bis keine Daten mehr im SRAM-Puffer REZ vorhanden sind.

35 Danach wird die Zeichenerkennungssoftware 224 aufgerufen. Bei der Ermittlung der Scanngeschwindigkeitswerte ist die Ermittlung der Steigungsmaße bei jedem der Weiß-schwarz- beziehungsweise Schwarz-weiß-Übergänge jeder Zeile nur dann erforderlich, wenn bei der Ermittlung der Scanngeschwindigkeitswerte über den Mittenabstand zweier benachbarter Zeichen Unsicherheiten in der Zeichenteilung bestehen. Dies kann der Fall sein, wenn Zeichen unterbrochen sind oder es keine Weißblücke zwischen den Zeichen gibt. Bei Proportionalsschriften sollte die Ermittlung der Scanngeschwindigkeiten über Steigungsmaße der Übergänge generell erfolgen, da bei diesen Schriften die Zeichenteilung nicht konstant ist. Eine exakte Geschwindigkeitskorrektur ist bei Proportionalsschriften nicht notwendig, da die Zeichen ohnehin unterschiedlich breit sind, und die ein Problem darstellt, welches durch die Zeichenerkennung gelöst werden muss.

40 Obwohl die Erfahrung anhand einer bevorzugten Ausführungsform beschrieben wurde, ist klar zu verstehen, dass sie hierauf nicht beschränkt ist. Die Einrichtung eines SRAM-Puffers ist nicht zwingend notwendig, sondern vorteilhaft. Selbstverständlich kann auch ein anderer Speicher verwendet werden, etwa ein DRAM oder ähnliches. Ebenso sind die in dem ausgeführten Beispiel angegebenen Zahlenwerte abänderbar. So können beispielsweise auch Scanner mit mehr als 64 Zeilen verwendet werden.

Patentansprüche

- 45 1. Verfahren zur Aufnahme von Bilddaten einer Zeichenzeile, wobei eine Zeichenzeile einer Vorlage zum Lesen von Bilddaten abgetastet und die Bilddaten für eine OCR-Erkennung vorbereitet werden, dadurch gekennzeichnet, dass die Zeichenteilung beruhend auf den Bilddaten bestimmt wird.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass die Zeichenteilung aus dem Abstand aufeinanderfolgender Zeichen bestimmt wird.
3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, dass zusätzlich die Übergangsschärfe an den Zeichenkanten bestimmt wird, wobei die Übergangsschärfe zur Überprüfung der aus dem Zeichenabstand ermittelten Zeichenteilung dient.
4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass die gelesenen Bilddaten komprimiert werden.
5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, dass die komprimierten Bilddaten als Grauwerte in einem SRAM einer Zeichenerkennungseinheit für Grauwert-Videodaten abgespeichert werden.
- 55 6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass es die Schritte
 - Komprimieren der eingelesenen Bilddaten und Speichern als Grauwerte in einem SRAM-Puffer
 - Abspeichern von Grauwertstützpunkten zum Bilden einer komprimierten Tapete
 - Bilden von Zeichenmittendenständen aufeinanderfolgender Zeichen
 - Bilden von Scanngeschwindigkeitswerten und Eintragen dieser Scanngeschwindigkeitswerte in die komprimierte Tapete
 - Dekomprimieren der komprimierten Tapeteumfasst.
7. Verfahren nach Anspruch 6, dadurch gekennzeichnet, dass beim Bilden der Grauwertstützpunkte die als Grauwerte abgespeicherten Bilddaten zeilenweise unter Bildung von Grauwertdifferenzen zu einem für jede Zeile ermittelten Weißwert untersucht werden, wobei bei Überschreiten einer vorbestimmten Grauwertdifferenz jeweils ein Grauwertstützpunkt in den SRAM-Puffer eingetragen wird.
- 60 8. Verfahren nach Anspruch 6 oder 7, dadurch gekennzeichnet, dass vor dem Bilden von Zeichenmittendenständen

DE 100 09 761 A 1

aufeinanderfolgender Zeichen Grauwerte unterhalb eines jeweiligen Schwellenwertes auf 0 gesetzt werden, wobei die Schwellenwerte jeweils zeilenweise abhängig von einem maximalen, in horizontaler Richtung ermittelten Grauwert gebildet werden.

9. Verfahren nach einem der Ansprüche 6 bis 8, dadurch gekennzeichnet, dass die Scanngeschwindigkeitswerte aus den Mittenabständen aufeinanderfolgender Zeichen mittels einer Mehrheitssuche bestimmt werden. 5
10. Verfahren nach Anspruch 9, dadurch gekennzeichnet, dass bei der Ermittlung der Scanngeschwindigkeiten die Steigungsmaße der Schwarz-weiß- beziehungsweise Weiß-schwarz-Übergänge der jeweiligen Zeichenteile in horizontaler Richtung ermittelt werden.
11. Bildaufnahmeverrichtung zur Aufnahme von Bilddaten einer Zeichenzeile mit einer Aufnahmeverrichtung zur Abtastung einer Vorlage, um die Bilddaten einer Zeichenzeile zu lesen, und einer Verarbeitungsvorrichtung, um unter Berücksichtigung einer tatsächlichen Zeichenzeile die Bilddaten für eine OCR-Erkennung vorzubereiten, um so verzerrungsfreie Bilddaten zu erhalten, 10 dadurch gekennzeichnet, dass die Verarbeitungsvorrichtung ausgestattet ist, um die Zeichenteilung beruhend auf den Bilddaten zu bestimmen.
12. Bildaufnahmeverrichtung nach Anspruch 11, dadurch gekennzeichnet, dass die Verarbeitungsvorrichtung ausgestattet ist, um den Abstand zwischen aufeinanderfolgenden Zeichen zu ermitteln und aus diesem Abstand die Zeichenteilung bestimmen. 15
13. Bildaufnahmeverrichtung nach Anspruch 11, dadurch gekennzeichnet, dass die Verarbeitungsvorrichtung die Übergangsschärfe an den Zeichenkanten bestimmt, wobei die Übergangsschärfe ein Maß für die Zeichenteilung ist. 20
14. Vorrichtung nach Anspruch 12, dadurch gekennzeichnet, dass die Verarbeitungsvorrichtung zusätzlich die Übergangsschärfe an den Zeichenkanten bestimmt, wobei die Übergangsschärfe zur Überprüfung der aus dem Zeichenabstand ermittelten Zeichenteilung dient.
15. Bildaufnahmeverrichtung nach Anspruch 11, dadurch gekennzeichnet, dass die Aufnahmeverrichtung ein handgeführter Scanner ist.
16. Bildaufnahmeverrichtung nach Anspruch 11, dadurch gekennzeichnet, dass die Vorrichtung einen Speicher aufweist, um die Bilddaten zwischenzuspeichern. 25

Hierzu 3 Seite(n) Zeichnungen

30

35

40

45

50

55

60

65

1
FIG

Zeile 28-> 0 1000000197051>330203000001005750002690768+ 010033262>

FIG 2

0100000197051>33020300001005750002690768+ 010033262>

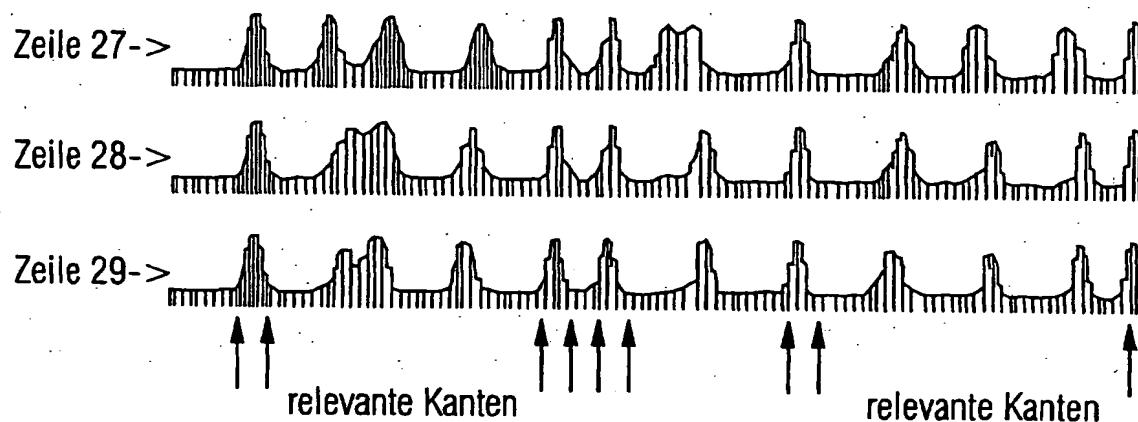
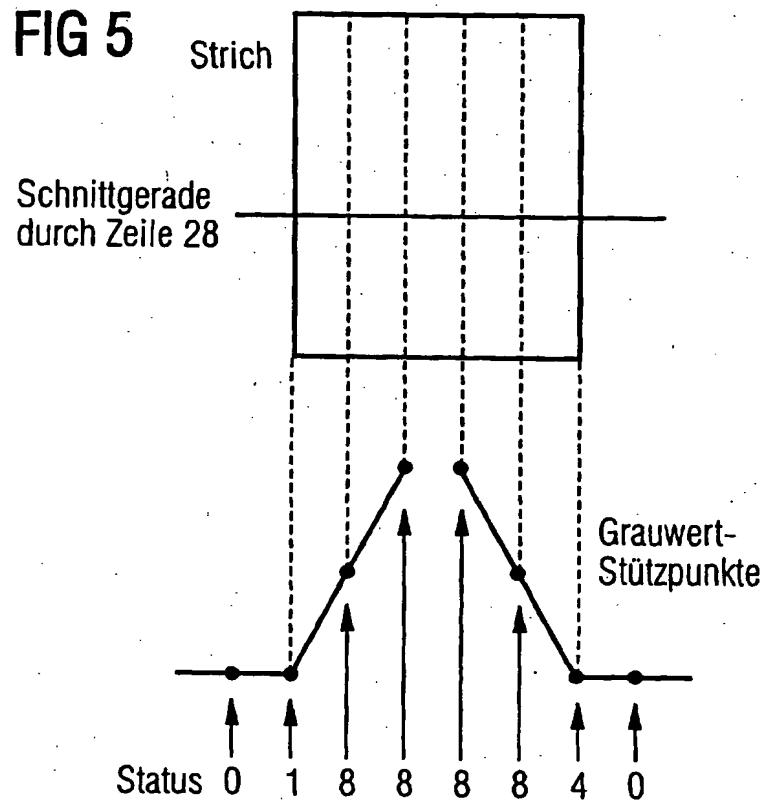
3

Zeichen 1 = '0' 1
Zeichen 2 = '0' 1
Zeichen 3 = '0' 1



FIG 4

Histogramm

**FIG 5**

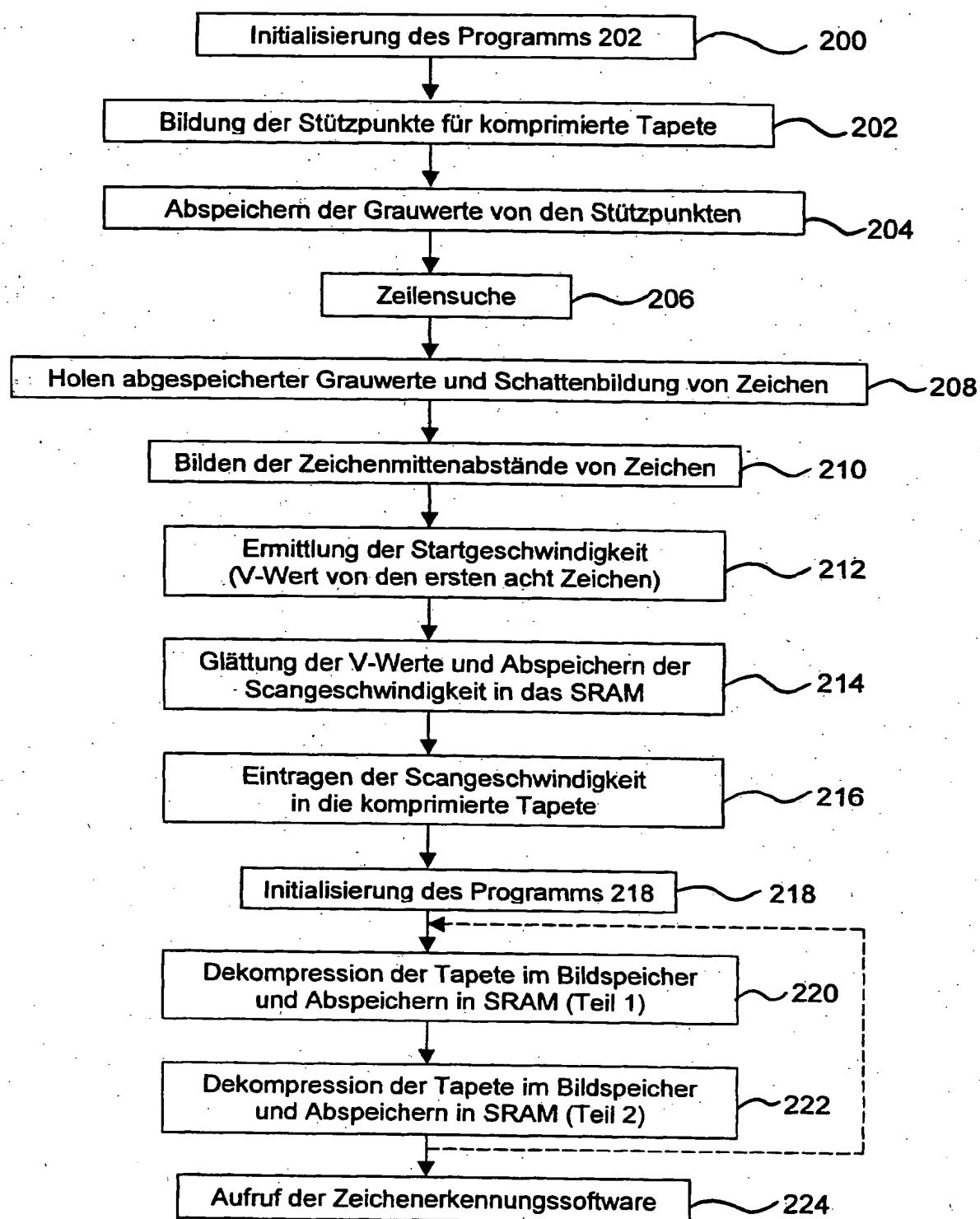


Fig. 6